METHOD AND APPARATUS FOR IDENTIFYING SEMANTIC STRUCTURES FROM TEXT

ABSTRACT OF THE DISCLOSURE

A method and apparatus for identifying a semantic structure from an input text forms at least two candidate semantic structures. A semantic score is determined for each candidate semantic structure based on the likelihood of the semantic structure. A syntactic score is also determined for each semantic structure based on the position of a word in the text and the position in the semantic structure of a semantic entity formed from the word. The syntactic score and the semantic score are combined to select a semantic structure for at least a portion of the text. In many embodiments, the semantic structure is built incrementally by building and scoring candidate structures for a portion of the text, pruning low scoring candidates, and adding additional semantic elements to the retained candidates.